



Training in Assessment of Preservice Teachers of Primary. How are the Opportunities of Learning Offered by Chilean Universities?

La formación de los futuros profesores de educación básica en evaluación. ¿Cómo son las oportunidades de aprendizaje que ofrecen las universidades chilenas?

Marianela Navarro Ciudad¹, Carla Förster², Ivonne Méndez Child³
& Lorena Meckes Gerard⁴

¹ Universidad de los Andes, Chile

² Universidad de Talca, Chile

³ Mineduc, Unidad de Currículum y Evaluación, Chile

⁴ Consejo Nacional de Educación, Chile

Abstract

It is not possible to think about the quality of education in a country without considering the training of teachers. Among the many skills that teachers in Chile need to develop, one of the weakest in this country is monitoring and assessing student learning. This represents a challenge for initial teacher training. The purpose of this study was to analyze the learning opportunities in assessment reported by 724 students training to be basic education teachers in the last year of their programs at 23 Chilean universities, who answered a survey about their preparation in terms of assessment of student learning in late 2014. The results show that little attention is given to the specificities of assessing the subjects they will teach, although there is a balance between theoretical and practical approaches. Given the sample studied, the availability of the instrument, and that the cohort analyzed was trained before the publication of the standards for initial teacher training (2012) and before the regulations for initial training were introduced under Law 20,903, this study constitutes a baseline to monitor possible changes as a consequence of these policies.

Keywords: learning assessment, assessment literacy, opportunities of learning, initial teacher training, pedagogy programs

Post to:

Marianela Navarro Ciudad
Monseñor Álvaro del Portillo 12455 Las Condes, Santiago, Chile
mnavarroc@uandes.cl

© 2021 PEL, <http://www.pensamientoeducativo.org> - <http://www.pel.cl>

ISSN:0719-0409 DDI:203.262, Santiago, Chile doi: 10.7764/PEL.58.2.2021.11

Resumen

No es posible pensar en la calidad de la educación de un país sin considerar la preparación de su profesorado. Entre las muchas competencias que los y las docentes necesitan desarrollar, la de monitorear y evaluar el aprendizaje de los escolares es una de las más descendidas en los profesores de Chile, lo que representa un desafío para su formación inicial. El propósito de este estudio fue analizar las oportunidades de aprendizaje en evaluación reportadas por 724 estudiantes de último año de carreras de pedagogía básica, de 23 universidades chilenas, quienes respondieron a un cuestionario sobre su preparación en el ámbito de la evaluación de aprendizajes a fines de 2014. Los resultados muestran un enfoque genérico para abordar este ámbito de la formación, con escaso énfasis en las particularidades que tiene la evaluación de aprendizajes de las disciplinas que deberán enseñar, aunque equilibrando aspectos teóricos y prácticos. Dada la muestra alcanzada, la disponibilidad del instrumento, y que la cohorte estudiada se formó en un periodo previo a la publicación de los estándares orientadores de la formación inicial en Chile (2012) y de las regulaciones introducidas por la Ley 20.903, este estudio constituye una línea de base para monitorear eventuales cambios derivados de estas políticas.

Palabras clave: evaluación de aprendizajes, alfabetización evaluativa, oportunidades de aprendizaje, formación inicial docente, carreras de pedagogía

Introduction

Training teachers in the skills they need to teach the new generations is perhaps the most important aspect to improve the quality of teaching and student learning (Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura, 2010). For this reason, specifically in Chile, a series of policies and regulations have been promoted that are intended to improve teacher training, with one of these being the definition of what they are expected to know and be able to do when they have completed their training, by means of the publication of standards for graduates of teaching programs (Ministerio de Educación, [Mineduc], 2012). These standards have been compulsory since 2016, because in order to achieve accreditation—which is also obligatory—faculties must comply with them in their graduate profiles and curricula. This is the case with the National Diagnostic Assessment (END, by the Spanish acronym), which has been applied mandatorily since 2017. This assessment measures the attainment of these standards one year before graduation, replacing the previous INICIA test for graduates, which was voluntary. It is also possible to point to the inclusion of a mentoring system, which helps the integration of novice teachers (Boerr, 2011), as well as the requirement to achieve a certain score on university admission tests in order to be able to study a teaching career, among other measures established under Law 20,903. Several of these measures had begun to be adopted by the education faculties of the universities belonging to the Council of Rectors of Chilean Universities (CRUCH) prior to the enactment of this law. For example, in 2011, they agreed to establish 500 points as the required minimum score on the University Admission Test (PSU); they had developed self-assessment processes, obtaining more years of accreditation than educational institutions not belonging to CRUCH; and they were actively involved in the development of the aforementioned standards.

Due to its huge impact on learning, the skill of assessing and monitoring student progress can be identified as one of those that are essential to develop in teacher training (Black, Harrison, Lee, Marshall, & Wiliam, 2004; Black & Wiliam, 2010; Deneen & Brown, 2016; Mertler & Campbell, 2015; Pastore & Andrade, 2019; Tejedor & García-Varcárcel, 2010; Torres & Cárdenas, 2010; Wiliam, Lee, Harrison, & Black, 2004). Indeed, assessment determines what and how students learn (Villagra, Sepúlveda, & Cerda, 2011), defines students' expectations of themselves, provides feedback on their performance (Black et al., 2004), and contributes to pedagogical decision-making on the part of teachers (Hamilton et al., 2009; Mandinach & Gummer, 2012).

In Chile, teachers' assessment practices have been identified by the National Teacher Performance Evaluation System as one of the weakest areas (Ministerio de Educación, 2020) and they have consequently become the focus of attention of both initial training policies and diagnoses of this training. Thus, the Guiding Standards for Elementary Education programs (Ministerio de Educación, 2012) outline clear requirements for the skills that graduates are expected to have developed in terms of assessing learning, including mastery of the specificities of assessment in each of the subjects. The Taskforce commissioned by the Ministry of Education to review the national learning assessment system also highlighted the need to improve initial teacher training in this area (Ministerio de Educación, 2015) and, more recently, specific studies on teacher training regarding learning assessment have been commissioned and carried out (Agencia de Calidad de la Educación, 2016; Gysling, 2017). It is evident that the focus on initial training as a lever to promote transformations in the capabilities of teaching staff is not limited to skills in learning assessment or to the Chilean context, as this has been a matter of concern in Chile and around the world in the last few decades (Darling-Hammond & Bransford, 2005; Floden, 2015; National Research Council, 2010; Ruffinelli, 2016).

Assessment literacy

Assessment literacy is the teaching skill to assess learning (Popham, 2009; Xu & Brown, 2016) and it is considered a professional requirement within the current accountability framework of public education (DeLuca, 2012; DeLuca, LaPointe-McEwan, & Luhanga, 2016; Popham, 2011). This skill is comprised of three dimensions: conceptual, praxeological, and socioemotional (Pastore & Andrade, 2019), which are very difficult to incorporate if they are not fostered in the initial training of preservice teachers (DeLuca, Chavez, & Cao, 2013). The first dimension refers to the teacher's knowledge of what assessment means, why to do it, how to assess, how to analyze the information collected, and how to report and communicate the results effectively to the various stakeholders. The second dimension entails evaluative practice and how to integrate assessment processes into other teaching practices in order to monitor and manage the teaching-learning process. This dimension has a component that is specific to each subject, which requires it to be combined with the didactic knowledge of the subject (Grossman, 1990; Magnusson, Krajcik, & Borko, 1999). The third dimension (socioemotional) takes into account that assessment is a social practice, which requires the management of aspects such as the attitude of the student and the way in which they face an evaluative process; ethical aspects, especially consequential validity, justice, and equity; the responsibility of the person who conducts the assessment; and the power and impact of the assessment on the student's commitment to learning and on the teacher-student relationship that will be created (Förster & Rojas-Barahona, 2017; Prieto & Contreras, 2008; Wiliam et al., 2004).

These generic dimensions are expressed in a particular way in each subject (Zolfaghari & Ahmadi, 2016), forming part of the pedagogical knowledge of the content (Schulman, 1987). This constitutes a greater requirement for training in the case of teachers of elementary education, since they teach more than one subject, particularly if one considers not only the conceptual dimension, but also the practical aspect (Grossman, 1990).

Problems identified in school teachers' assessment practices

At the international level, research on learning assessment practices in schools has shown that there is a predominant approach in which assessment is conceived as being dissociated from the pedagogical process, and where memorization of content is emphasized (Celman, 2005; Goubeaud, 2010; Organisation for Economic Co-operation and Development, [OECD], 2005; Prieto & Contreras, 2008; Sanmartí, 2007). Another problem identified is the use of assessment and specifically grading as an instrument of control (Torres & Cárdenas, 2010). In addition, in Latin America it has been observed that feedback practices are more evaluative than descriptive and, therefore, they provide little guidance (Ravela, 2009). A negative effect has also been reported on low-performing students (Black et al., 2004), since there is a tendency to highlight errors without taking advantage of them to enhance learning (Torres & Cárdenas, 2010). Similarly, assessments that value and promote simple cognitive skills are preferred (Black et al., 2004; Prieto & Contreras, 2008), and explicit grading criteria are uncommon (Ravela, Leymoníe, Viñas, & Haretche, 2014).

In a comparative study that examined the assessment instruments used by elementary school teachers in several Latin American countries, Ravela et al. (2014) showed that, in Chile, written tests were the instrument most frequently used with clear predominance of closed, multiple-choice questions, lacking context and focused on retrieving information. Meanwhile, the results of the National Teacher Performance Evaluation System show that there are low or very low levels of achievement of the indicators associated with learning assessment: the use of error for learning (10%), feedback for students (22%), analysis and use of assessment results (17%), design of assessment instruments (34%), and coherence between assessment tasks and learning objectives (44%) (Ministerio de Educación, 2019a).

Given this background, it is clear that these weaknesses in the assessment practices of teachers in Chile pose a challenge for the training of new generations of teachers.

Learning assessment and initial teacher training

It can be hypothesized that the weaknesses identified among in-service teachers are to some extent due to deficient university training in the skill of learning assessment. In this respect, assessment has historically been an area that has been neglected in teaching programs, with little research to support the training of preservice teachers in this skill (DeLuca, 2012).

Data from the END test indicate that student teachers of elementary education reach an achievement level of 55% for the standard on assessment, while in the area of reflection on teaching practice—which includes key skills in the assessment process such as analysis and decision making (Hamilton et al., 2009; Mandinach & Gummer, 2012; Ministerio de Educación, 2018)—they attain a level of 47% (Ministerio de Educación, 2019b). These results provide little information to indicate how satisfactory the level of knowledge and skills is in each case, since it is not clear whether the measurement allows direct comparison of the data for the different standards assessed and because no cutoff score is established to define the minimum level that is considered acceptable.

A descriptive study conducted by the Education Quality Agency (2016) using a sample of 14 elementary education programs in Chile provides more information on initial teacher training (ITT) in the area of learning assessment. The study shows that most of these programs only have one course on assessment in the entire study plan. Although there is content on assessment within the courses on teaching and, in some cases, in the practicum, these are not priority topics and there is little connection between these subjects. The study also suggests that the only course on assessment offered in teacher training programs in Chile is excessively theoretical or lacking

in depth in terms of the process of gathering evidence, in the analysis of results, and in pedagogical decision-making. According to the study, these deficiencies are mainly due to the lack of opportunities throughout the training programs to apply the theoretical designs learned repeatedly and progressively.

This is the case not only in Chile. In a systematic review of 100 international studies, Xu and Brown (2016) state that many teacher training programs offer only a one-semester course on assessment that provides a general introduction, which is often theory-based and disconnected from actual classroom assessment practices.

Although the assessment course in ITT in Chile generally involves an emphasis on the educational function of assessment, in practice the focus is on the design of instruments and not on providing descriptive feedback to students, which is a fundamental element of formative assessment (Gysling, 2017). Similarly, the assessment instruments that are designed are not applied, so they are not contrasted empirically, and the information on learning that can be obtained by applying them is not analyzed. According to Gysling (2017), this is exacerbated because professional practices are far from being a space in which students can apply what is taught theoretically at university.

In summary, according to the information examined, the opportunities offered to student teachers of elementary education to learn and develop their skills to assess learning are not sufficient to train competent and literate teachers to do this (Coombs, DeLuca, LaPointe-McEwan, & Chalas, 2018).

Opportunities of learning for preservice teachers

Opportunities of learning (OTLs) refer to teaching inputs and processes that influence the attainment of expected learning. The opportunity to learn can be linked to the content on which work is done, to the amount of time students spend on activities intended for their learning, and also to the quality of teaching (Elliot & Barlett, 2016; Kurz, 2011). Among the elements that these authors consider to comprise OTLs, time and content have a moderate effect with respect to student learning, while teaching quality has a moderate to large effect (Kurz, 2011).

Opportunities to learn can occur at the level of the prescribed or intended curriculum, which then need to be crystalized in the implemented curriculum. In order to research the OTLs in schools, there have been frequent studies that examine the recommended curriculum and study texts, and in university education, there has been analysis of curricula and study programs. This analysis can shed light on the dimensions of time and content (e.g., number of courses, number of explicit learning objectives focused on certain content) in the intended curriculum. This is the approach of analyses of initial teacher training curricula in Chile to study the learning opportunities they provide (Cofré et al., 2010; Sotomayor, Parodi, Coloma, Ibáñez, & Cavada 2011; Varas et al., 2008).

Studying and measuring effective OTLs requires going beyond the prescribed curriculum and addressing the implemented curriculum. This presents various challenges, including simultaneously encompassing the dimensions of content coverage, teaching time, and teaching quality, and managing to do this reliably (Elliot & Barlett, 2016). The methodological approaches used to study this include logs and surveys of teachers, student reports of their learning experiences in interviews and questionnaires, observation and coding of classes, and analysis of student work (Floden, 2002; Klette & Hammerness, 2016). Even though students' self-reporting of their formative experience via surveys may have limitations due to being mediated by memory, among other factors, it is encouraging that student reports show positive correlations with learning outcomes (Floden, 2002). Indeed, in the field of teacher education, a study conducted by König, Ligtvoeta, Klemenza, and Rothland (2017) in 37 Austrian and German university programs found that both theoretical and practical OTLs, measured using student questionnaires, were predictive of outcomes on a test assessing their pedagogical knowledge in the areas of planning, adaptation of teaching to diversity, and also in learning assessment, which is highly relevant to our study.

In Chile, the OTLs of preservice teachers have also been studied through questionnaires to inquire about their formative experience in the subject of Language (Sotomayor-Echeñique et al., 2013) and to compare the opportunities to learn certain teaching practices on different courses of the curriculum (Müller, Álamos, Meckes, Sanyal, & Cox, 2018).

Objective and questions guiding the study

In Chile, the standards for ITT in elementary education (Ministerio de Educación, 2012) define the expectations regarding what a teaching graduate should know and do in the area of learning assessment, among other things. However, they do not provide guidance for teacher training institutions in terms of implementation, nor do they specify which OTLs should be provided in order to develop these skills. This study focuses specifically on investigating the OTLs required to meet the standards on learning assessment by means of a questionnaire for student teachers of elementary education.

To summarize, research on the processes to train preservice teachers to develop their evaluative skills is important because this is an area in which the performance of practicing teachers is deficient and due to the large impact it has on children's learning (Black et al., 2004; Black & William, 2010; Tejedor & García-Varcárcel, 2010; Torres & Cárdenas, 2010). In spite of its importance, there are still few studies on this subject in Chile. The objective of this paper is to analyze the OTLs for learning assessment reported by prospective teachers of elementary education during their university training.

We also aim to answer the following questions: How do opportunities to learn about assessment compare with the overall OTLs of initial teacher training in elementary education at each training institution? Does a theoretical or practical approach predominate in these learning opportunities? Do they emphasize general knowledge and skills related to assessment or the particularities of assessment in each subject? Are there differences between these OTLs in institutions that belong to CRUCH and those that do not? All of these questions are addressed using the self-reporting of student teachers of elementary education in Chile.

According to the literature reviewed and the research questions, we pose the following hypotheses: OTLs in assessment are lower than the overall OTLs of initial teacher training in each of the institutions. OTLs in assessment provided by teacher training institutions in elementary education in Chile show a theoretical emphasis rather than practical OTLs in the area of learning assessment. OTLs in assessment provided by teacher training institutions in elementary education in Chile show a general emphasis rather than a specific focus on assessment of the subjects that teachers will teach. And, finally, the OTLs in assessment reported by student teachers differ depending on whether their institution is a member of the CRUCH group of universities.

The relevance of the study lies in the opportunity to produce empirical evidence that describes the opportunities to learn about assessment reported by student teachers of elementary education prior to the introduction of national policies in ITT and before this was a focus of these policies. The results reported here thus constitute a baseline to monitor the evolution of the OTLs provided by ITT in learning assessment.

Methodology

Design

The methodology is based on a descriptive and comparative design, which is intended to characterize the OTLs in ITT in the area of learning assessment according to students' reports. The study is a secondary analysis of data from the FONDEF D11|1109 project *Elaboración, validación y aplicación de instrumentos de diagnóstico de oportunidades de aprendizaje para el logro de los estándares nacionales en la formación de profesores de educación básica* (Preparation, validation, and application of diagnostic instruments of learning opportunities for the attainment of national standards in the training of teachers of elementary education).

Population and sample

The population comprised 47 Chilean universities that offered elementary education programs with face-to-face and daytime courses. All of these universities were invited to participate in the study using census methodology. Through a web platform, each interested institution had to register the students who were in their last year of training. In order to ensure the representativeness of each institution, only those universities that completed the survey with a response rate of at least 60% of the students enrolled were included in the sample. Thus, 23 institutions distributed throughout Chile were included, with a total of 724 students, who responded to the instrument in late 2014. Because of the voluntary nature of the survey, the sampling was not intended to be representative of the population; however, we obtained a sample with similar characteristics to the population in terms of geographical distribution of the institutions and whether or not they were members of CRUCh.

Forty-nine percent of the institutions were covered, of which 45.7% ($n=11$) were universities that were members of CRUCh and 54.3% ($n=12$) were not. This distribution was very similar to that nationwide, where 42.6% of ITT institutions were CRUCh members and 57.4% were not. In descriptive terms, statistically significant differences were reported in favour of CRUCh institutions in terms of PSU admission scores ($t(722)=5.90, p<0.001$) and years of accreditation ($t(722)=8.97, p<0.001$) (Table 1). The teacher training institutions were distributed throughout Chile, with 11% in the north, 57% in the center, and 32% in the south of the country, which is very similar to the distribution of the universe of institutions in the country (11%, 54%, and 35%, respectively).

Table 1
Characterization of the institutions that comprise the study sample

Institution	n	Area of the country	Type of institution	Description	Average PSU score and SD	Average number of years of accreditation and SD
1	90	Center	CRUCh	These are state and non-state public universities with a presence in all regions of Chile. They have a single admissions system, the right to receive direct contributions from the state and university credit. The rectors hold ordinary meetings in which they discuss pedagogical and management issues (Council of Rectors ¹).	490.44 (192.46)*	5.03 (0.87)*
2	13	South				
3	18	South				
4	39	South				
5	22	North				
6	37	North				
7	15	South				
8	26	Center				
9	15	North				
10	21	South				
11	35	Center				
Subtotal	331					
12	19	Center-South	Non-CRUCh	These are the private universities in the country. They may or may not be affiliated to the single admissions system for Chilean universities of the Council of Rectors.	401,60 (212.26)*	4.12 (1.67)*
13	94	Center-South				
14	33	Center-South				
15	22	Center				
16	51	Center-South				
17	18	Center				
18	22	Center				
19	37	Center				
20	22	Center				
21	27	Center				
22	16	Center				
23	32	Center				
Subtotal	393					
Total	724					

* $p < 0.001$

Source: Prepared by the authors.

1. <https://www.consejodirectores.cl/universidades>

Instrument

The instrument is a questionnaire on OTLs that are conducive to attaining general pedagogical and disciplinary knowledge for teaching in elementary education, as defined in the Guiding Standards for ITT (Ministerio de Educación, 2012). Questionnaires to investigate learning opportunities in initial teacher training have also been used in other countries (e.g., Klette & Hammerness, 2016).

The questionnaire consisted of 211 items that addressed OTLs for general pedagogical knowledge (about the school curriculum, assessment, learning and development, pedagogical interaction, teaching design, and planning), disciplinary knowledge, and specific didactic knowledge in the areas of Mathematics, Language, Natural Sciences, History, and Social Sciences, using a 4-point numerical response scale, where 1 represented None or very few OTLs and 4 represented Extensive or many OTLs. The students responded to the questionnaire via a website.

In order to construct the items, we designed a four-quadrant matrix, where (1) general pedagogical knowledge and (2) pedagogical content knowledge were cross-referenced with (3) theoretical and (4) practical learning opportunities. The first two conceptual categories are based on those of Schulman (1987): general pedagogical knowledge (GPK) and pedagogical content knowledge (PCK). Distinguishing between formative and summative assessment could be an example of general knowledge about learning assessment, while identifying the most appropriate type of assessment for a given scientific thinking skill would be an example of pedagogical content knowledge, more closely linked to the teaching of the subject. The other two categories respond to the type of approach to teacher training: either theoretical or discursive, based on training models in which it is conceived that one learns to teach based on a solid conceptual background: one must know about teaching in order to teach (Korthagen, 2010). This theoretical approach can be shown by OTLs that are characterized by expository lectures or classes on assessment, while the practical classes require the students to be exposed to practical testing experiences or application in schools, on the premise that it is necessary to teach in order to know about teaching (Korthagen & Kessels, 1999). From this perspective, the future teacher is involved in activities such as, for example, analyzing authentic work by schoolchildren, or designing instruments and applying them to students. The cross-referencing and description of these categories is shown in Table 2, along with examples of the items.

Table 2

Description and examples of items according to category of pedagogical knowledge and type of OTLs

	General pedagogical knowledge on assessment	Pedagogical knowledge of the content on assessment	Description of sub-index
By means of a theoretical or discursive approach	A discursive and general approach to knowledge on assessment, such as studying or reading research on the relationship between motivation and learning assessment, attending lectures on different types of assessment depending on their purpose (2 items).	A discursive and domain-dependent approach to knowledge on assessment, such as reading or studying habitual errors or preconceptions of schoolchildren in a specific subject area, reading about the results of assessments of schoolchildren's scientific thinking skills (5 items).	Theoretical OTLs sub-index: involves a discursive approach to knowledge on assessment (7 items).

By putting it into practice or approaching practice	OTLs that are characterized by the application of general knowledge about assessment, such as observing a teacher in a video using a strategy to provide feedback in classes, identifying whether it is descriptive or evaluative, analyzing the numerical data of a learning assessment (6 items).	OTLs that are characterized by an applied approach of domain-dependent knowledge on assessment, such as examining authentic work by students to understand their preconceptions in a specific subject area, designing an assessment task to assess understanding of a fundamental idea in science, or conducting experiments with a student or group of students to demonstrate their preconceptions in the same subject area (15 items).	Practical OTLs sub-index: involves an applied approach to knowledge on assessment (21 items).
Description of sub-index	General OTLs sub-index: involves the mastery of general knowledge on assessment (8 items).	Domain-dependent OTLs sub-index: involves the mastery of specific knowledge about assessment in the subjects (20 items).	

Source: Prepared by the authors.

We obtained evidence of the validity of the instrument through expert judgment and the validity analysis of the construct. The team of experts working directly on the project was the first source to analyze the items. In a second stage, the instrument was reviewed by national and international experts (Robert Floden, Karen Hammerness, Kirsty Klette, and Elizabeth Davis) who were external to the project, specialists in learning opportunities in ITT, and who had participated in similar initiatives. The analysis criteria were: the relevance and alignment of the items and the instrument as a whole with the findings in research on teacher training and with the standards for elementary education graduates (Ministerio de Educación, 2012), and also the number and usefulness of the questions asked in order to form consistent indices.

Immediately after completing the pilot questionnaire, we conducted five group interviews with the students to collect information about the characteristics and contents of the instrument and the platform, to identify the main difficulties involved in applying it, and to gather recommendations to improve the process.

For the construct validity, we carried out an exploratory and confirmatory factor analysis to select the items for each dimension. We considered the items that had factorial loads greater than 0.3 in the theoretical factor and, subsequently, we selected the groups of questions that achieved fit indices as close as possible to an RMSEA index of less than 0.05 and a CFI of greater than 0.9 (Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco, 2014). The questions selected to construct the indices were therefore those whose confirmatory factor analysis models converged and produced the best indicators of fit and which showed adequate internal consistency (Cronbach's alpha) for both the instrument and the dimensions (Nunnally & Bernstein, 1995).

In this study, we worked with the 28 items referring to OTLs in the area of learning assessment during initial teacher training, which were grouped into a general index and into four sub-indices obtained from different combinations of items:

- Total index of OTLs in assessment: entails the combination of the 28 items mentioned above. It includes questions on general pedagogical knowledge in the area of assessment and questions on pedagogical knowledge of the subject to assess this learning, where the OTLs have had both a theoretical-discursive and practical approach ($\alpha = 0.94$).
- Theoretical-practical sub-indices: these two sub-indices were created from items of theoretical OTLs in assessment (7 items; $\alpha = 0.75$) and practical OTLs (21 items; $\alpha = 0.93$). The sub-index of theoretical OTLs refers to a discursive approach to knowledge in assessment, such as reading, studying, or listening to lectures on specific learning assessment topics. The sub-index of practical OTLs in assessment refers to teaching where preservice teachers have exercised skills related to learning assessment, such as by designing assessment instruments, applying them, and analyzing their results.
- Sub-indices of general pedagogical knowledge or subject-specific knowledge of assessment: these two sub-indices involve items related to OTLs regarding general aspects of assessment (8 items; $\alpha = 0.92$) and OTLs for focused assessment of language, mathematics, natural sciences and history, geography, and social sciences (20 items; $\alpha = 0.93$). The sub-index of OTLs for general aspects of assessment entails content such as types of assessments according to their intentionality, agent, or reference, or types of learning feedback. The sub-index referring to OTLs to assess learning specifically in the subjects that a teacher of elementary education has to teach involves content such as providing feedback to students on their oral or written production, designing assessments that take into account the errors that students commonly make in science, formulating indicators to assess civic attitudes and values, or designing strategies to assess mathematical reasoning.
- Finally, we considered the overall average of the OTLs in initial teacher training, which summarizes the students' responses to the full questionnaire on OTLs (211 items) covering the other areas of training (disciplinary knowledge in the four subjects mentioned, on the school curriculum, on learning and child development, on school culture, etc.).

The robustness of the resulting indices should be noted, since they are the result of an exhaustive and rigorous validation process, are composed of a set of items, and show internal consistency values that are higher than those expected for this type of instruments.

Data analysis

We first obtained averages for each student teacher in terms of the general OTLs in assessment of learning and for each of the sub-indices shown in Table 2.

The analyses are presented in accordance with the questions that guided the study. So, to answer the research question, 'How do opportunities to learn about assessment compare with the overall OTLs of initial teacher training in elementary education at each training institution?', we compared the result in the total index of the OTLs in assessment for each institution with the overall average of the OTLs for complete initial training, as reported by the students (corresponding to the average of the 211 items that make up the full instrument). In order to make this comparison, the score obtained for the OTLs in learning assessment was standardized, considering the mean and standard deviation of the overall OTLs of ITT by institution. It should be noted that these institutional parameters were used because the formative experiences in a given field do not occur in isolation; they involve various characteristics with which student teachers coexist within their respective institutions (Calixto & Herrera, 2010). Consequently, we considered that the best way to establish the degree to which students experienced more or fewer learning opportunities in the areas of learning assessment was to contrast their opinions with those on the training received at their institution in the other areas, avoiding aggregating and directly comparing the OTLs reported by students from different institutions. The OTLs in assessment expressed for each student in Z scores (considering the institutional mean and standard deviation of the overall OTLs of ITT) were converted into a categorical variable, considering the cutoff points defined by

Cohen (1988, 1992). The levels of OTLs in assessment that were considered were: much lower than the overall average OTLs reported for the institution (≤ -0.7), lower than the average (> -0.7 and < -0.3), similar to the average (≥ -0.3 and ≤ 0.3), more positive than the average (>0.3 and < 0.7), and much more positive than the average (≥ 0.7). In addition, possible differences between OTLs in assessment between institutions belonging to the CRUCh group and those not belonging to the CRUCh group (hereafter “non-CRUCh”) were analyzed using a χ^2 test, with a statistical significance of 5%.

To answer the second question, ‘Does a theoretical or practical approach predominate in these learning opportunities?’, we compared the average obtained in the indices of theoretical and practical OTLs in assessment in each institution. These differences were also analyzed according to whether the institution belonged to CRUCh or was non-CRUCh. Similarly, in order to address the third question, ‘Do they emphasize general knowledge and skills related to assessment or the particularities of assessment in each subject?’, we compared the average obtained in the indices of OTLs in assessment of general pedagogical knowledge and OTLs in assessment of specific knowledge in the subjects for each institution. As in the previous analysis, we examined these differences depending on whether the institution was a member of CRUCh or not.

The size of the differences was also analyzed according to Cohen’s standardized differences for all the analyses.

Results

The results are organized below in accordance with the research questions. With respect to the first question, ‘How do opportunities to learn about assessment compare with the overall OTLs of initial teacher training in elementary education at each training institution?’, we found that, according to the reports from the student teachers, in the majority of the institutions (15 of 23), most of the students report that there are fewer OTLs in assessment compared with what they state overall for the OTLs that consider all areas of initial teacher training. This is also evident in the full bar in Figure 1, which shows the heterogeneity of the results. Thus, for almost half of the total students in the sample (48%), the reported learning opportunities are lower than the overall average of learning opportunities in their initial training. On the other hand, 36% of the sample report that OTLs in assessment are more positive than the overall average of OTLs in initial training. For the remaining 16% of the students, no differences are observed in their views of OTLs for assessment and for the formative experience as a whole.

It is interesting to note that the responses within each institution are also heterogeneous, since there is a group of students in all of them who report fewer OTLs involving assessment and another group reporting more OTLs.

Analyzing the institutions specifically and according to Cohen’s standardized differences method, we can observe that there is no institution in which 50% or more of the students report more OTLs in assessment compared with the overall OTLs in initial training. Those that come closest are institutions 6 (49%) and 21 (48%), the former being a CRUCh institution and the latter non-CRUCh. There are institutions in which the students report OTLs in assessment being fewer than in others. For example, in institutions 3 (CRUCh), 12, and 23 (non-CRUCh), 60% or more of the students reported fewer OTLs in assessment than OTLs in the other areas of their training as a whole. There are institutions—10 (CRUCh) and 14 (non-CRUCh)—where the percentage of students who report OTLs in assessment as being more positive than the formative experience as a whole equals the percentage of students at the same institution who report the opposite.

Finally, when contrasting the OTLs in assessment versus the overall OTLs and differentiating by the type of institution (CRUCh and non-CRUCh), we observed no differences ($\chi^2(16, N=5) = 20.0, p=0.220$) (Figure 1).

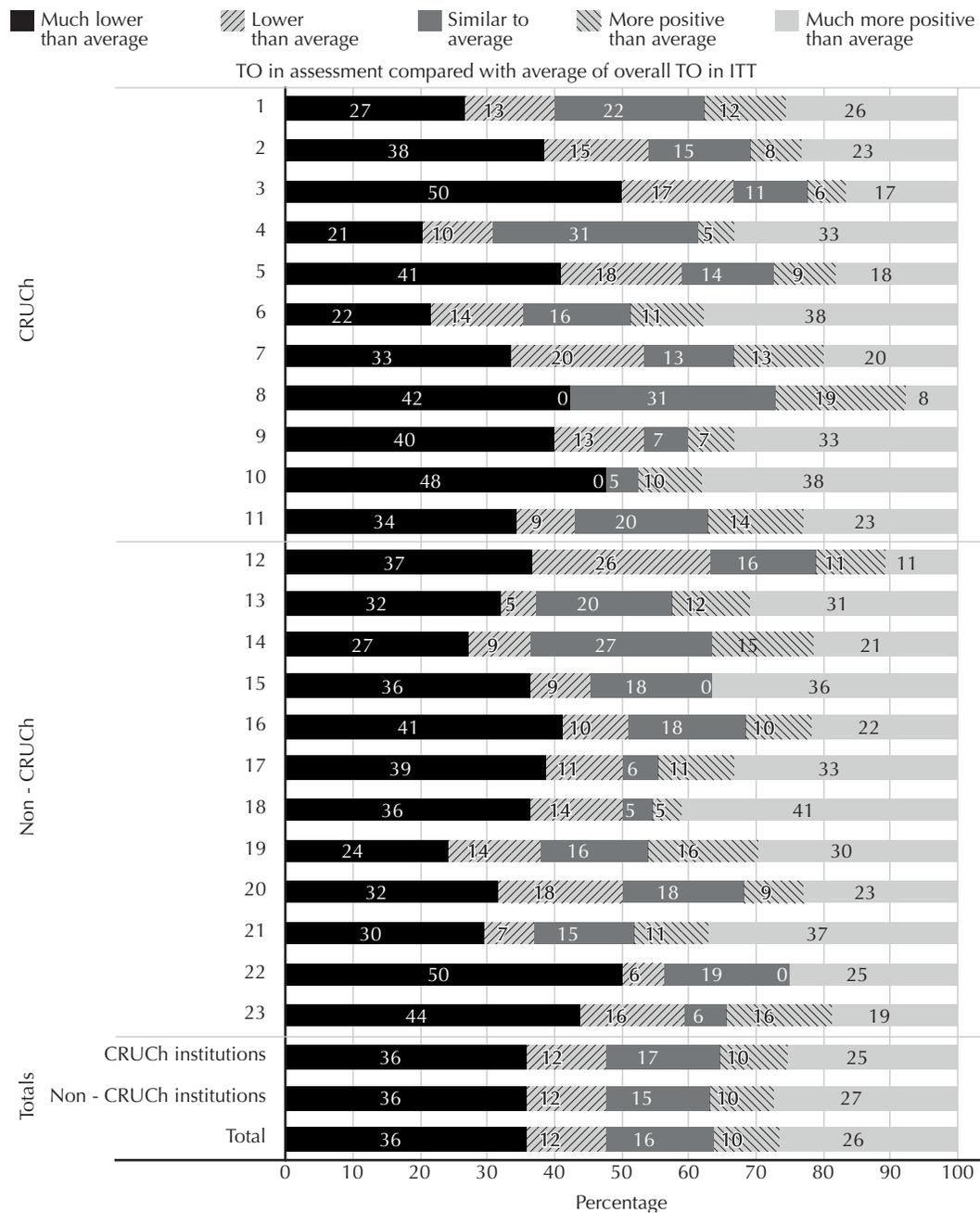
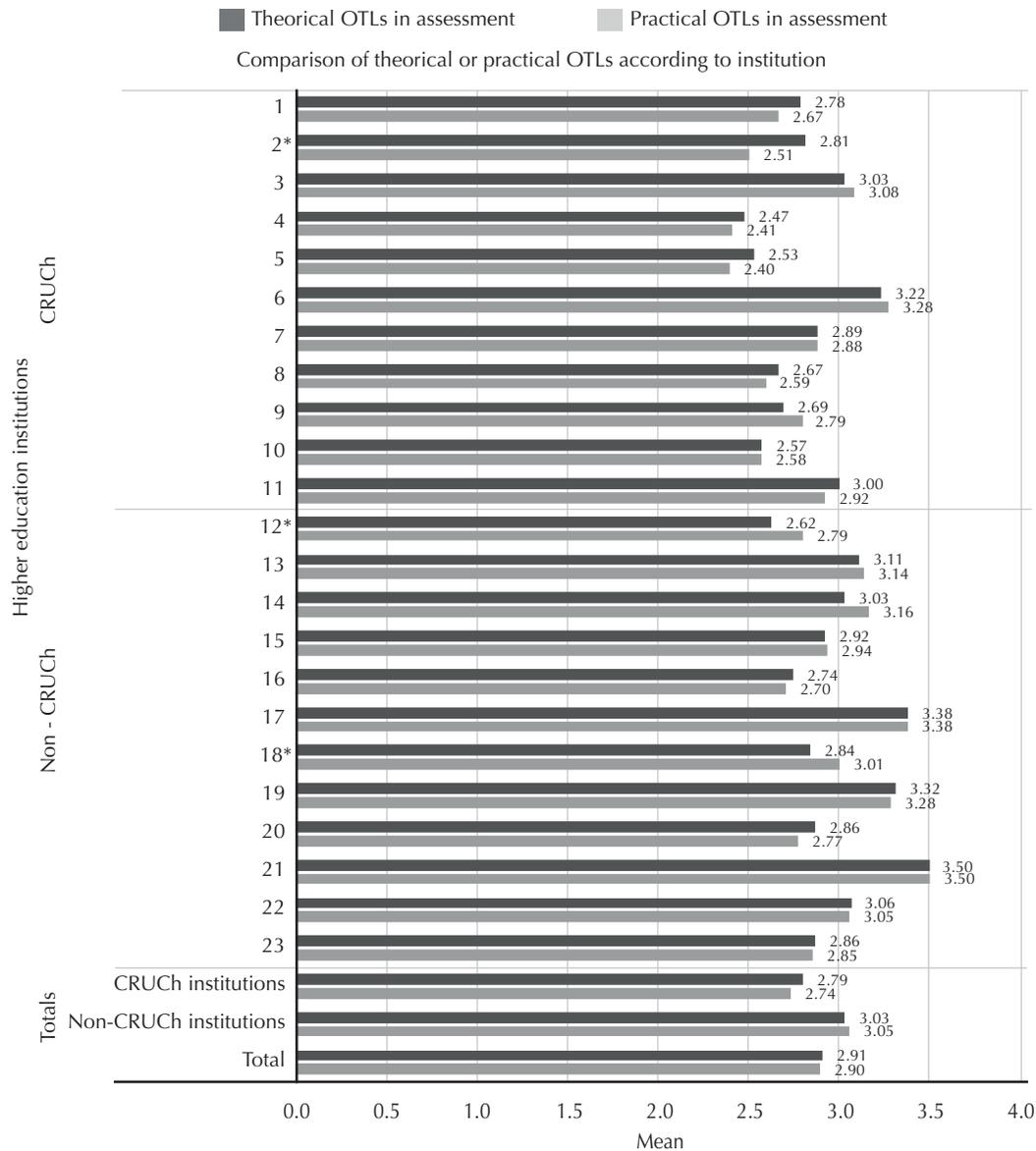


Figure 1. Level of OTLs in assessment in higher education institutions compared with the overall average OTLs in ITT.

Source: Prepared by the authors.

Regarding to the second question, ‘Does a theoretical or practical approach predominate in learning opportunities related to assessment?’, we find that in most of the institutions (20 of 23) the students report a balance between both types of opportunities ($d=0.03$) (Figure 2). However, we can see differences in some specific institutions, such as in institution 2 (CRUCh), where there is an emphasis on theoretical OTLs, while in institutions 12 and 18 (non-CRUCh) the emphasis is on practical experiences. At the aggregate level, we observe no differences between CRUCh ($d=0.10$) and non-CRUCh ($d=-0.04$) institutions according to Cohen’s standardized differences method.



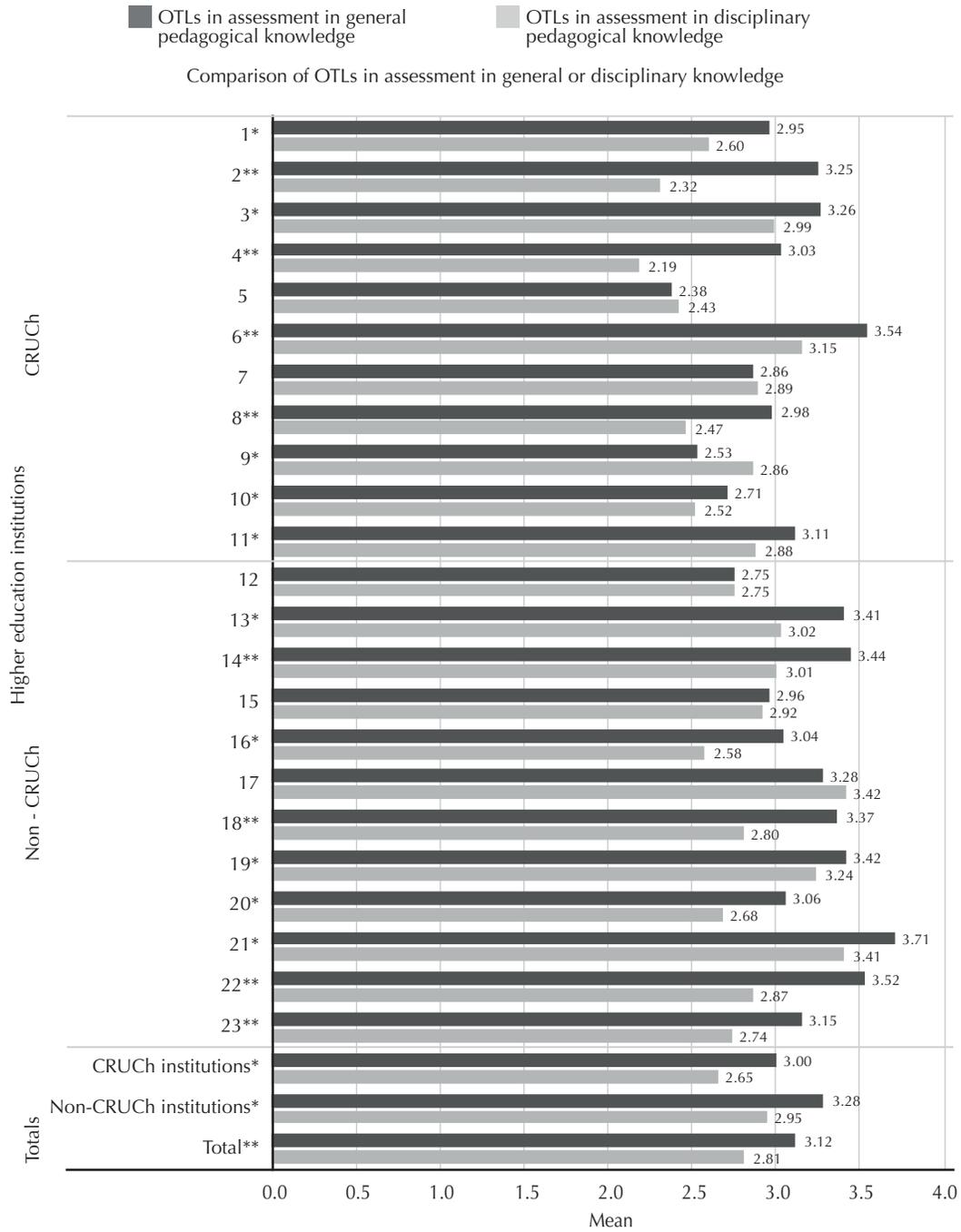
* $d > 0.3$

Figure 2. Comparison of OTLs in theoretical or practical assessment according to higher education institution.

Source: Prepared by the authors.

In relation to the third question, ‘Do they emphasize general knowledge and skills related to assessment or the particularities of assessment in each subject?’, the students report more learning opportunities for general pedagogical knowledge about assessment than OTLs about the particularities of assessment in each subject ($d=0.92$) (Figure 3). More specifically, in 18 of the 23 institutions examined, we did find moderate differences in 10 of them ($d > 0.3$), while the differences are more marked in the remaining eight ($d > 0.7$).

At the aggregate level, in CRUCh ($d=0.51$) and non-CRUCh ($d=0.53$) institutions, we see the same trend as described above, with moderate differences according to Cohen’s standardized differences method.



* $d > 0.3$

** $d > 0.7$

Figure 3. Comparison between OTLs in assessment in general or disciplinary knowledge according to higher

Source: Prepared by the authors.

Discussion and Conclusions

The objective of this study was to analyze OTLs in assessment reported by preservice teachers of elementary education in Chile during their university training, with the aim of characterizing whether their approach was theoretical or practical, and whether it focused on generic elements of learning assessment or if it was specific to the various subjects. We also sought to compare the reports of students in training institutions depending on whether or not they belonged to the CRUCH group of universities.

The results showed that, for most of the initial teacher training institutions, the OTLs in assessment reported by prospective elementary education teachers in Chile during their university training generally turned out to be fewer than those reported for other areas of ITT. This result is consistent with findings described by the literature reviewed both in Chile and in other countries, where weaknesses in initial teacher training have been identified in this area. For example, it is a fact that most of the curricula for elementary education programs have only one introductory-style course on assessment that is disconnected from classroom practices and courses on disciplinary didactics (Agencia de la Calidad de la Educación, 2016; Gysling, 2017; Xu & Brown, 2016).

Another aim of our study was to examine the applied component of evaluative practice in training (Gysling, 2017; Pastore & Andrade, 2019; Villagra et al., 2011). With regard to the question ‘Does a theoretical or practical approach predominate in these opportunities of learning?’, the students’ reports indicate there is a balance between experiences of practical application and experiences focused on the conceptual or theoretical dimension. This is promising, as it is essential for student teachers to have the opportunities to practice the skills needed to assess learning, such as preparing assessment tasks, analyzing evidence of learning, making judgments, providing feedback to students, and making pedagogical decisions.

This result differs from what Xu and Brown (2016) stated in their systematic review, where they pointed out that, in the area of assessment, teacher training tends to have a theoretical emphasis and little connection to classroom practices. Likewise, Gysling (2017) suggests that preservice teachers seem not to have experience of applying the instruments they design with real students. Therefore, the opportunity is lost for them to judge the clarity of instructions, identify types of possible responses, determine ways to define scores, discriminate between different levels of performance, provide feedback to students, or state possible pedagogical decisions. The apparent discrepancy between these studies and our results may lie in the way that the pedagogy of practice is conceived in teacher education. In our study, we include among the practical OTLs all of those that can be conceptualized as approaches to practice, also considering simulation experiences in controlled settings (Grossman, Hammerness, & McDonalds, 2009), and not only experiences at practice centers or in authentic school contexts. We therefore include experiences of creating instruments, discriminating between different feedback alternatives, and analyzing real assessment results, and more as practical learning opportunities. Under this perspective of practical training, the initial hypothesis (that OTLs in assessment offered by elementary education teacher training institutions in Chile show a clear theoretical emphasis rather than practical OTLs) is not confirmed in our study. However, addressing the practical component, as we understand it, requires design and follow-up in a progressive model from approaches to practice in controlled contexts to practical experiences in actual school settings.

With respect to mastering assessment from a disciplinary perspective rather than a general one, there is a recognizable need for teachers to have assessment knowledge that is specific to the subject they teach, particularly in association with the assessment of skills such as writing and speaking in Language; problem solving and reasoning skills in Mathematics and Physics; and experimentation and understanding of figures and tables in Science, to name just a few (Zolfaghari & Ahmadi, 2016). That said, do the learning opportunities for prospective teachers emphasize the particularities of assessment in each subject or do they focus on general knowledge and

skills in assessment? According to the students' reports, a general approach is predominant rather than addressing the specificities of assessment in the subjects they will teach. It is likely that this emphasis stems from the scant integration of assessment into the teaching courses for the subjects (Agencia de la Calidad de la Educación, 2016) and also from conceiving assessment as primarily summative, focused on certification and not on the learning process (Stiggins, 2004). On the other hand, the results could also be explained by the emphasis that assessment courses in ITT for elementary or primary education have historically had, focusing on preparing preservice teachers to develop decontextualized instruments based on a more psychometric approach (particularly tests) (Deneen & Brown, 2016; DeLuca & Klinger, 2010), without considering that each subject has skills and knowledge that require a particular proficiency on the part of the teacher (Grossman, 1990; Sanmartí & Alimenti, 2004; Tacochi & Fernández, 2014). On the other hand, when assessment focuses on its formative purpose, it is impossible to dissociate it from its specific teaching and content, since evaluative activities become part of the learning process and not a separate achievement (William et al., 2004). It has been reported that the assessment strategy depends on the teacher's conceptions of teaching-learning and can affect students' representations of the subject and their learning (Hofstein, Mamlok-Naaman, & Rosenberg, 2006; Sanmartí & Alimenti, 2004).

If under the Disciplinary Standards for Elementary Education (Ministerio de Educación, 2012) it is considered that prospective teachers at this level should master the particularities of teaching and assessment of at least four subjects, these results establish a baseline that shows a gap and a clear challenge to improve initial training, especially if it is considered that the findings of the most recent research do not show any great changes in training, raising a red flag regarding how much training on assessment is done in didactic-disciplinary courses (Agencia de la Calidad de la Educación, 2016; Gysling, 2017). The hypothesis that indicates that the OTLs in assessment provided by elementary education teacher training institutions in Chile show a general emphasis rather than a specific emphasis on the assessment of the subjects that teachers will teach is therefore confirmed.

Previous studies in Chile have shown differences in the reports of graduates or students on their training, depending on the selectivity of the institution where they studied (Ruffinelli, 2014). Considering that, during the period in which the questionnaire was applied, the CRUCH institutions agreed to increase their admission requirements by voluntarily establishing a minimum admission score, thus becoming more selective, we could have expected to find differences in the reports of students from institutions that belong to this group and those who studied at non-CRUCH institutions. However, this was not the case. We found no significant differences between the two groups of institutions for any of the analyses carried out. This is consistent with the study by the Education Quality Agency, which shows that the patterns of training on learning assessment are rather generalized and do not differ between the types of training institutions (Agencia de Calidad de la Educación, 2016).

As a conclusion of the study, the OTLs reported by the students show there is a disconnection in training on assessment between general and specific disciplinary aspects, although there is apparently a balance between theoretical and practical aspects. This disconnection may highlight the low coverage of evaluative elements in the didactic courses of the subjects, which raises a red flag for urgent consideration in ITT. Developing the skill to assess learning does not depend on a single course in degree programs, but each course of the curriculum should be a model of assessment, in addition to connecting this evaluative knowledge in the didactic and practical courses. Based on the results of this study and the findings of other research conducted in Chile, new questions arise, to which those who train preservice teachers of elementary education should pay attention: Is an assessment course sufficient to develop the skill to assess learning? Is assessment sufficiently integrated into the didactic courses for the subjects and in professional practices? How can a synergy be achieved in the OTLs in assessment for the different subjects? And what impact are the public policy guidelines of ITT having on training learning assessment?

Finally, studying and monitoring OTLs through student responses to questionnaires such as the one used in this research is a practice that is reported internationally (Klette & Hammerness, 2016). The instrument designed for this study is a validated tool that is available for use² and which is potentially useful for institutions to track the OTLs they are offering, based on information collected quickly and at a low cost.

Limitations and Future Research

One of the limitations of this study is the fact that the numbers of items in the different indices were not equivalent. This is due to the fact that the full questionnaire included questions on the overall pedagogical and disciplinary knowledge required by a preservice teacher and it was not designed exclusively to measure the OTLs in assessment provided by teacher training institutions. A second limiting aspect is measuring the OTLs offered in each training program based on student reports, especially considering the variations within the institutions. As described, the reports are heterogeneous in each institution, since there is a group of students for whom the OTLs in the specific area of learning assessment are significantly fewer than in the Global training experience, while for another group, there is no difference from the rest of the training or, on the contrary, they feel there are more OTLs for learning assessment. Like other similar studies, in this study we opted to consider the average of the responses or the predominant pattern of reports within each institution. However, it would be interesting to examine whether the differences in the opinions of the various groups of students when reporting on their OTLs are linked to their characteristics³ or whether they reflect actual differences, as would be the case if they attended different sections of the same program.

In order to enhance the measurement and analysis of the OTLs in assessment during initial teacher training and to validate the use of this instrument, which is based on students' self-reporting, it could be studied whether these reports show a relationship with results on measurements of their knowledge or performance in learning assessment, as was done in the study by König et al. (2017).

A second extension of this research would be to study the characteristics of institutions that offer more or fewer OTLs in assessment from a qualitative perspective, examining factors such as teaching strategies in ITT in the area of learning assessment, both in general and specific didactic courses.

It would also be interesting to apply the questionnaire again with novice teachers of elementary education in order to verify whether their reports on the OTLs in assessment change when compared with the demands of professional practice in schools. This is due to the fact that, according to the study by the Education Quality Agency (2016), the professional practices of prospective teachers are mediated by supervisors and, therefore, novice teachers have a better idea of the reality of teaching once they start working professionally in schools.

Finally, the most important possible extension of this study would be to remeasure the OTLs in assessment based on the validated instrument presented here. This new research would make it possible to monitor the impact of ITT policies in the area of learning assessment, considering the results presented here as a baseline obtained at the time when these policies were in the early phase of implementation.

Funding: Project FONDEF D11|1109 Preparation, validation, and application of diagnostic instruments of learning opportunities for the achievement of national standards in training teachers for basic education.

2. On request from the authors.

3. See for example the study by Cohen and Berlin (2020), which reveals a relationship between the personality characteristics of ITT students and their reports of learning opportunities in degree courses.

The original paper was received on March 11th, 2020
 The reviewed paper was received on September 9th, 2020
 The paper was accepted on November 10th, 2020

References

- Agencia de Calidad de la Educación. (2016). *Estudio sobre formación inicial docente en evaluación educacional*. Retrieved from http://archivos.agenciaeducacion.cl/liderazgo-motivacion-lectora/Resumen_Ejecutivo_FID_en_Evaluacion_educacional.pdf
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working Inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan*, 86(1), 8-21. <https://doi.org/10.1177/003172170408600105>
- Black, P. & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 92(1), 81-90. <https://doi.org/10.1177/003172171009200119>
- Boerr, I. (2011). Inserción profesional de los docentes, desarrollo de la mentoría en Chile. *Nuevos Paradigmas de las Ciencias Sociales Latinoamericanas*, 2(3), 81-86. Recuperado de <https://nuevosparadigmas.ilae.edu.co/index.php/IlaeOjs/article/view/209>
- Calixto, R. & Herrera, L. (2010). Estudio sobre las percepciones y la educación ambiental. *Tiempo de Educar*, 11(22), 227-249.
- Celman, S. (2005). ¿Es posible mejorar la evaluación y transformarla en una herramienta de conocimiento? In A. R. W. Camilloni, S. Celman, E. Litwin, & M. C. Palou de Maté (Comp.), *La evaluación de los aprendizajes en el debate didáctico contemporáneo* (pp. 35-66). Buenos Aires, Argentina: Paidós.
- Cofré, H., Camacho, J., Galaz, A., Jiménez, J., Santibáñez, D., & Vergara, C. (2010). La educación científica en Chile: debilidades de la enseñanza y futuros desafíos de la educación de profesores de ciencia. *Estudios Pedagógicos*, 36(2), 279-293. <https://doi.org/10.4067/s0718-07052010000200016>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. <http://doi.org/10.1037/0033-2909.112.1.155>
- Cohen, J. & Berlin R. (2020). What Constitutes an “Opportunity to Learn” in Teacher Preparation? *Journal of Teacher Education*, 71(4), 434-448. <https://doi.org/10.1177/0022487119879893>
- Coombs, A., DeLuca, C., LaPointe-McEwan, D., & Chalas, A. (2018). Changing approaches to classroom assessment: An empirical study across teacher career stages. *Teaching and Teacher Education*, 71, 134-144. <https://doi.org/10.1016/j.tate.2017.12.010>
- Darling-Hammond, L. & Bransford, J. (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco, CA: Jossey-Bass.
- DeLuca, C. (2012). Preparing teachers for the age of accountability: toward a framework for assessment education. *Teacher Education Yearbook XXI: A Special Issue of Action in Teacher Education*, 34(5-6), 576-591. <https://doi.org/10.1080/01626620.2012.730347>
- DeLuca, C. & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates’ learning. *Assessment in Education: Principles, Policy y Practice*, 17(4), 419-438. <https://doi.org/10.1080/0969594X.2010.516643>
- DeLuca, C., Chavez, T., & Cao, C. (2013). Establishing a foundation for valid teacher judgement on student learning: The role of pre-service assessment education. *Assessment in Education: Principles, Policy and Practice*, 20(1), 107-126. <https://doi.org/10.1080/0969594X.2012.668870>
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Teacher assessment literacy: a review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28, 251-272. <https://doi.org/10.1007/s11092-015-9233-6>
- Deneen, C. & Brown, G. T. L. (2016). The impact of conceptions of assessment-on-assessment literacy in a teacher education program. *Cogent Education*, 3(1), 1-14. <https://doi.org/10.1080/2331186X.2016.1225380>
- Elliot, S. N. & Barlett, B. J. (2016). *Opportunity to Learn*. <https://doi.org/10.1093/oxfordhb/9780199935291.013.70>

- Floden, R. (2002). The Measurement of Opportunity to Learn. In A. C. Porter and A. Gamoran (Eds.), *Methodological Advances in Cross-National Surveys of Educational Achievement*, (pp. 231-266). Washington, DC: National Academy Press.
- Floden, R. (2015). Learning what research says about teacher preparation. In M. J. Feuer, A.I. Berman, & R. C. Atkinson (Eds.), *Past as prologue: The national academy of education at 50. members reflect* (pp. 279-284). Washington, DC: National Academy of Education.
- Förster, C. E. & Rojas-Barahona, C. (2017). Aprendizaje y evaluación: lo que no se evalúa, no se aprende. In C. Förster (Ed.), *El poder de la evaluación en aula. Mejores decisiones para promover aprendizajes* (1st. ed., pp.43-74). Santiago, Chile: Ediciones UC.
- Goubeaud, K. (2010). How is Science Learning Assessed at the Postsecondary Level? Assessment and Grading Practices in College Biology, Chemistry and Physics. *Journal of Science Education and Technology*, 19, 237-245. <https://doi.org/10.1007/s10956-009-9196-9>
- Grossman, P. L. (1990). *The making of a teacher: Teacher and teacher education*. New York, NY: Teachers College Press.
- Grossman, P., Hammerness, K., & McDonald, M. (2009) Redefining teaching, re-imagining teacher education, *Teachers and Teaching*, 15(2), 273-289. <https://doi.org/10.1080/13540600902875340>
- Gysling, J. (2017). *La evaluación: ¿dispositivo para promover el aprendizaje de todos o para seleccionar?: La formación de profesores en evaluación en Chile* (Doctoral Dissertation, Universidad Diego Portales). Retrieved from <https://openaccess.leidenuniv.nl/handle/1887/46245>
- Hamilton, L., Halverson, R., Jackson, S. S., Mandinach, E., Supovitz, J. A., Wayman, J. C., ... & Steele, J. L. (2009). *Using Student Achievement Data to Support Instructional Decision Making*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from https://repository.upenn.edu/cgi/viewcontent.cgi?article=1298&context=gse_pubs
- Hofstein, A., Mamlok-Naaman, R., & Rosenberg, O. (2006). Varying instructional methods and assessment of students in high school chemistry. In M. McMahon, P. Simmons, R. Sommers, D. De Baets, & F. Crawley (Eds.), *Assessment in science* (pp. 139–148). Arlington, TX: NSTA.
- Klette, K. & Hammerness, K. (2016). Conceptual Framework for Analyzing Qualities in Teacher Education: Looking at Features of Teacher Education from an International Perspective. *Acta Didactica Norge*, 10(2), 26-52. <https://doi.org/10.5617/adno.2646>
- Korthagen, F. (2010). How teacher education can make a difference. *Journal of Education for Teaching: International research and pedagogy*, 36(4), 407-423. <https://doi.org/10.1080/02607476.2010.513854>
- Korthagen, F. & J. Kessels (1999). Linking theory and practice: Changing the pedagogy of teacher education. *Educational Researcher*, 28(4), 4-17. <https://doi.org/10.3102/0013189X028004004>
- König, J., Ligtoet, R., Klemenza, D., & Rothland, M. (2017). Effects of opportunities to learn in teacher preparation on future teachers' general pedagogical knowledge: Analyzing program characteristics and outcomes. *Studies in Educational Evaluation*, 53, 122–133. <https://doi.org/10.1016/j.stueduc.2017.03.001>
- Kurz, A. (2011). Access to what should be taught and will be tested: Students' opportunity to learn the intended curriculum. In S. N. Elliott, R. J. Kettler, P. A. Beddow, & A. Kurz (Eds.), *The handbook of accessible achievement tests for all students: Bridging the gaps between research, practice, and policy* (pp. 99-130). New York, NY: Springer.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El Análisis Factorial Exploratorio de los Ítems: una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169. <https://doi.org/10.6018/analesps.30.3.199361>
- Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge: The construct and its implications for science education* (pp. 95- 132). Boston, MA: Kluwer.
- Mandinach, E. & Gummer, E. (2012). *Navigating the Landscape of Data Literacy: It IS Complex*. Retrieved from https://www.wested.org/online_pubs/resource1304.pdf
- Mertler, C. & Campbell, C. (2015, april 15). *Measuring Teachers' Knowledge and Application of Classroom Assessment Concepts: Development of the Assessment Literacy Inventory*. Paper presented at the annual meeting of the American Educational Research Association, Montréal, Quebec, Canada. Retrieved from <http://files.eric.ed.gov/fulltext/ED490355.pdf>

- Ministerio de Educación. (2012). *Estándares orientadores para egresados de carreras de pedagogía en educación básica* (2nd ed.). Santiago, Chile: LOM Ediciones.
- Ministerio de Educación. (2015). *Informe Equipo de Tarea para la Revisión del SIMCE*. Retrieved from <https://www.mineduc.cl/wp-content/uploads/sites/19/2015/11/Informe-Equipo-de-Tarea-Revisi%C3%B3n-Simce.pdf>
- Ministerio de Educación. (2018). *Política para el fortalecimiento de la evaluación en aula*. Retrieved from <https://bibliotecadigital.mineduc.cl/handle/20.500.12365/2255>
- Ministerio de Educación. (2019a). *Resultados nacionales Evaluación Docente 2018*. Retrieved from <https://bibliotecadigital.mineduc.cl/handle/20.500.12365/14547>
- Ministerio de Educación. (2019b). *Resultados nacionales Evaluación Nacional Diagnóstica de la Formación Inicial Docente 2018*. Retrieved from <https://bibliotecadigital.mineduc.cl/handle/20.500.12365/4660>
- Ministerio de Educación. (2020). *Resultados Nacionales Evaluación Docente 2019*. Retrieved from <https://bibliotecadigital.mineduc.cl/handle/20.500.12365/14861>
- Müller, M., Álamos, P., Meckes, L., Sanyal, A., & Cox, P. (2018). Percepción de estudiantes de pedagogía en relación a las oportunidades para el desarrollo de prácticas generativas en su formación. *Estudios Pedagógicos*, 42(4), 145-163.
- National Research Council. (2010). *Preparing teachers: Building evidence for sound policy*. Washington, DC: The National Academies Press.
- Nunnally, J. C. & Bernstein, I. H. (1995). *Teoría Psicométrica*. Mexico City, Mexico: McGraw Hill.
- Organisation for Economic Co-operation and Development. (2005). *Formative Assessment: Improving Learning in Secondary Classrooms*. Retrieved from <http://www.oecd.org/edu/cei/35661078.pdf>
- Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura. (2010). *2021 metas educativas la educación que queremos para la generación de los bicentenarios*. Retrieved from <http://www.redage.org/sites/default/files/adjuntos/metast2021-2.pdf>
- Pastore & Andrade, (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, 84, 128-138. <https://doi.org/10.1016/j.tate.2019.05.003>
- Popham, W. J. (2009). Assessment Literacy for Teachers: Faddish or Fundamental? *Theory Into Practice*, 48(1), 4-11. <https://doi.org/10.1080/00405840802577536>
- Popham, W. J. (2011). Assessment Literacy Overlooked: A Teacher educator's Confession. *The Teacher Educator*, 46(4), 265-273. <https://doi.org/10.1080/08878730.2011.605048>
- Prieto, M. & Contreras, G. (2008). Las concepciones que orientan las prácticas evaluativas de los profesores: un problema a develar. *Estudios Pedagógicos*, 34(2), 245-262. <https://doi.org/10.4067/S0718-07052008000200015>
- Ravela, P. (2009). Consignas, devoluciones y calificaciones: el problema de la evaluación de las aulas de educación primaria en América Latina. *Páginas de educación*, 2, 49-89. Retrieved from <http://paginasdeeducacion.ucu.edu.uy/inicio/item/download/14.html>
- Ravela, P., Leymoní, J., Viñas, J., & Haretche, C. (2014). La evaluación en las aulas de secundaria básica en cuatro países de América Latina. *Propuesta educativa*, 41(1), 20-45. Retrieved from <http://propuestaeducativa.flasco.org.ar/wp-content/uploads/2019/12/41-dossier-ravelayotros.pdf>
- Ruffinelli, A. (2014). Dificultades de la iniciación docente. ¿Iguales para todos? *Estudios Pedagógicos (Valdivia)*, 40(1), 229-242. <https://doi.org/10.4067/S0718-07052014000100014>
- Ruffinelli, A. (2016). Ley de desarrollo profesional docente en Chile: de la precarización sistemática a los logros, avances y desafíos pendientes para la profesionalización. *Estudios Pedagógicos (Valdivia)*, 42(4), 261-279. <https://doi.org/10.4067/S0718-07052016000500015>
- Sanmartí, N. (2007). *10 ideas clave Evaluar para Aprender*. Barcelona, Spain: Graó.
- Sanmartí, N. & Alimenti, G. (2004). La evaluación refleja el modelo didáctico: análisis de actividades de evaluación planteadas en clases de química. *Educación Química*, 15(2), 120-128. <https://doi.org/10.22201/fq.18708404e.2004.2.66198>
- Schulman, L. (1987). Knowledge and Teaching: Foundations of the New Reform. *Harvard Educational Review*, 57(1), 1-23. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Stiggins, R. J. (2004). *Classroom assessment for student learning: Doing it right—using it well*. New York, NY: Pearson Education.

- Sotomayor, C., Parodi, G., Coloma, C. J., Ibáñez, R., & Cavada, P. (2011). *La formación inicial de docentes de educación general básica en Chile: ¿Qué se espera que aprendan los futuros profesores en el área de Lenguaje y Comunicación?* Retrieved from <http://ciperchile.cl/wp-content/uploads/Formacion-Inicial-profesoreslenguaje-1.pdf>
- Sotomayor-Echenique, C., Coloma-Tirapegui, C. J., Parodi-Sweis, G., Ibañez-Orellana, R., Cavada-Hrepich, P., & Gysling-Caselli, J. (2013). Percepción de los estudiantes de pedagogía sobre su formación inicial. *Magis. Revista Internacional de Investigación en Educación*, 5(11), 375-392. <https://doi.org/10.11144/Javeriana.m5-11.peps>
- Tacochi, M. & Fernández, C. (2014). Knowledge of assessment: an important component in the PCK of chemistry teachers. *Problems of education in the 21st Century*, 62, 124-147. Retrieved from http://www.scientiasocialis.lt/pec/node/files/pdf/vol62/124-147.Tacoshi_Vol.62.pdf
- Tejedor, F. & García-Varcárcel, A. (2010). Evaluación de desempeño docente. *Revista Española de Pedagogía*, (68)247, 453-459. Retrieved from <https://revistadepedagogia.org/lxviii/no-247/evaluacion-del-desempeno-docente/101400010164/>
- Torres, M. & Cárdenas, E. (2010). ¿Qué y cómo se ha investigado sobre la evaluación de los aprendizajes en los últimos cinco años? Estado del arte de las investigaciones (2005–2010). *Enunciación*, 15(1), 141-156. <https://doi.org/10.14483/22486798.3109>
- Varas, L., Felmer, P., Gálvez, G., Lewin, R., Martínez, C., Navarro, S., Ortiz, A., & Schwarze, G. (2008). Oportunidades de preparación para enseñar matemáticas de futuros profesores de educación general básica en Chile. *Calidad en la Educación*, 29, 64-88. <https://doi.org/10.31619/caledu.n29.188>
- Villagra, C., Sepúlveda, S., & Cerda, P. (2011). Los desafíos de la evaluación en la formación inicial docente. *Revista Investigaciones en Educación*, 11(2), 35-44. Retrieved from <http://revistas.ufro.cl/ojs/index.php/educacion/article/view/1038>
- William, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education*, 11(1), 49-65. <https://doi.org/10.1080/0969594042000208994>
- Xu, H. & Brown, G. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149-162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Zolfaghari, F. & Ahmadi, A. (2016). Assessment literacy components across subject matters, *Cogent Education*, 3(1), 1-16. <https://doi.org/10.1080/2331186X.2016.1252561>

Appendix 1

Items that comprise the OTLs in assessment

Indices	Items
<p style="text-align: center;"><i>OTLs from a practical approach</i></p> <p style="text-align: center;"><i>OTLs for general pedagogical knowledge in assessment</i></p>	<p>To what extent did the program give you opportunities to examine different instruments and strategies for learning assessment (e.g., studying the construction of tests, guidelines for observation, assessment criteria, their coherence with the objectives to be assessed, among other aspects)?</p>
	<p>To what extent did the program give you opportunities to observe evaluative practices that promote student learning (e.g., describing how a teacher gives feedback on learning, or uses error as a source of learning)?</p>
	<p>To what extent did the program give you opportunities to design assessment strategies and instruments, according to teaching methodologies, learning objectives and assessment purposes (e.g., diagnostic, formative, summative instruments)?</p>
	<p>To what extent did the program give you opportunities to apply assessment strategies or instruments to a student or course?</p>
	<p>With regard to the use of the results of assessment, to what extent did the program give you opportunities to communicate assessment results and provide feedback for the students' learning process and promote their self-regulation?</p>
<p style="text-align: center;"><i>OTLs from a theoretical or discursive approach</i></p>	<p>With regard to learning assessment, to what extent did the program give you opportunities to study types of assessment (e.g., formal and informal, conceptual, procedural, attitudinal, diagnostic, formative, summative)?</p>
	<p>With regard to the use of assessment results, to what extent did the program give you opportunities to study procedures to analyze the results of learning assessment (e.g., statistical analysis, rubrics, process analysis)?</p>
<p style="text-align: center;"><i>OTLs for pedagogical knowledge on assessment of subjects</i></p>	<p>To what extent did the program give you opportunities to use assessment results to diagnose students' potential and difficulties in listening comprehension?</p>
	<p>To what extent did the program give you opportunities to give useful and timely feedback to students about their oral production to help them improve it?</p>
	<p>To what extent did the program give you opportunities to take decisions based on students' lexical, grammatical and spelling performance?</p>

	<p>To what extent did the program give you opportunities to design and/or <i>apply various assessment instruments</i> that would make it possible to observe the development of reading comprehension levels and strategies?</p>
	<p>To what extent did the program give you opportunities to <i>design assessments</i> considering the habitual errors of the students in learning various mathematical topics (e.g., considering habitual errors to prepare distractors in multiple choice questions).</p>
	<p>To what extent did the program give you opportunities to learn to <i>assess</i> the achievement of mathematical processes <i>during a class</i> (e.g., problem-solving, mathematical reasoning, argumentation and justification)?</p>
<p><i>OTLs</i> from a practical approach</p>	<p>To what extent did the program give you opportunities to <i>construct assessment questions</i> using knowledge about the errors that students make when applying adding or subtraction algorithms?</p>
	<p>To what extent did the program give you opportunities to <i>design strategies to assess</i> the skills related to discovering regularities and expressing them in formulas?</p>
<p><i>OTLs</i> for pedagogical knowledge on assessment of subjects</p>	<p>To what extent did the program give you opportunities to <i>prepare relevant assessment instruments</i> to monitor learning in the social sciences (such as concept maps, research papers, debate, among others)?</p>
	<p>To what extent did the program give you opportunities to <i>design an instrument</i> to assess an historic study (e.g., rubric, scale of assessment, etc.)?</p>
	<p>To what extent did the program give you opportunities to <i>assess civic attitudes and values</i> in specific situations (e.g., debates, organization of a project, collaborative work, etc.)?</p>
<p><i>OTLs</i> from a theoretical or discursive approach</p>	<p>To what extent did the program give you opportunities to <i>have expository classes, read, study, or discuss</i> strategies to assess both the process and the written production of students considering different criteria (e.g., clarity, precision of ideas, use of lexis, and resources used)?</p>
	<p>To what extent did the program give you opportunities to <i>have expository classes, read, study, or discuss</i> strategies to assess the achievement of mathematical learning (e.g. levels of complexity, elaboration of distractors)?</p>
	<p>To what extent did the program give you opportunities to <i>have expository classes, read, study, or discuss</i> strategies to assess skills such as conjecturing and demonstrating rules of pattern formation (levels of complexity, preparation of distractors)?</p>
	<p>To what extent did the program give you opportunities to <i>have expository classes, read, study, or discuss</i> about assessment instruments to evaluate learning of natural sciences (e.g., KPSI, previous ideas, rubric for correction of laboratory reports)?</p>